Random Proportional Weibull Hazard Model for Predicting Deterioration of Educational Facilities

Kiyoyuki KAITO[1], Kengo OBAMA[2] and Kiyoshi KOBAYASHI[3]

1. Associate Professor, Frontier Research Center, Osaka University, Japan
2. Master Course Student, Graduate School of Urban Management, Kyoto University, Japan
3. Professor, Graduate School of Management, Kyoto University, Japan

Abstract: In asset management of infrastructures, predicting deterioration of structures is one of an essential technique. Especially this study focuses on educational facilities from among social infrastructures. However, as for such kind of the facilities that consist of a large number of structural components, in order to estimate their deterioration process with high accuracy, the heterogeneity of individual components has to be considered because each component possesses different material characteristics and designs and is in unique service under various environmental conditions. This paper addresses the random proportional Weibull hazard model that is possible to consider the heterogeneity of individual components as probability fluctuation in the hazard rate. Specifically, the time-dependent hazard rate is formulated by the Weibull hazard model. The heterogeneity of the hazard rates across the individual characteristics of components is explained by the random proportional Weibull hazard model in which the hazard rates are subject to Gamma distribution. This study also presents the estimation techniques for the proposed model based on the maximum likelihood estimation method. In addition, through an empirical study employing visual inspection data for an actual university facility, with deterioration prediction for components, the validity of the method is verified.

Keywords: Educational Facilities, Asset Management, Random Proportional Weibull Hazard Model, Deterioration Prediction, Heterogeneity

1. INTRODUCTION

In the same way as civil infrastructures, Japanese educational facilities have been built continuously from the period of high economic growth. In general the expected lifetime of educational facilities is about 30 years, which is short in comparison to civil infrastructures. In fact, it has been pointed out that for educational facilities that were built in the early stages of the period of high economic growth, their repair and reconstruction costs began to surface around 2005, sooner than for civil infrastructures. It can easily be deduced that these costs weigh down the management of educational facilities, and it is absolutely necessary to develop asset management to support various decision makings regarding the planning of repair and reconstruction strategies.

In the asset management of educational facilities, lifecycle cost is an important evaluation index that determines repair and reconstruction strategies. In addition, deterioration prediction results are reflected in the evaluation of lifecycle costs, and so the establishment of deterioration prediction technique is also an important issue. In general, deterioration prediction methods can be roughly classified as: 1. physical deterioration prediction methods based on the mechanical deterioration mechanisms of structural components and 2. statistical deterioration methods based on past inspection data. However, for educational facilities, repairs and reconstructions are sometimes carried out based not only on physical deterioration but also on the users' usability and visual factors (aesthetics). Therefore, when attempting to carry out deterioration predictions for educational facilities, it is preferable to employ a statistical deterioration prediction method.

Statistical deterioration prediction methods are methods that take vast amounts of deterioration information and model the regularities behind deterioration processes. In recent years there has been a remarkable accumulation of research into deterioration models using hazard functions [1-4]. Hazard models are distinctive because in characterizing the deterioration process of each facility they respond to the structural characteristics of the facility and environmental conditions to give individual hazard rates. However, as a hazard rate is given deterministically, the deterioration process for facilities that have the same structural characteristics and environmental conditions will be identical. Regarding this point, even when structural characteristics and environmental conditions are the same, it is more natural to consider that the deterioration process will differ for each facility. Therefore, in order to carry out a more exhaustive deterioration prediction, it is necessary to develop a deterioration prediction method that takes into account the heterogeneity of the deterioration process for individual components.

With an awareness of the above issues, in this study the authors propose a random proportional hazard model that expresses the heterogeneity of the individual deterioration processes of facilities as a hazard rate probability distribution. Specifically, a random proportional Weibull hazard model in which the hazard rates are subject to Gamma distribution is developed. Below, in section **2** the basic concepts of this study are consolidated, in section **3** the random proportional Weibull hazard model and its estimation methods are explained, and in section **4** as a empirical study of application, a university facility is taken up and some analysis carried out based on its visual inspection data.

2. BASIC CONCEPTS OF THIS RESEAECH

2.1. Current state of educational facilities

Educational facilities includes public, national and private schools (elementary, junior high and senior high schools), research institutions, day cares, kindergartens, universities, research institutions, museums, art galleries, libraries and community facilities, as well as social education facilities, lifelong learning facilities and cultural and community facilities fall into this category. In the same way as other social infrastructures, as a part of economic policy in the postwar era, the educational facilities in Japan were constructed as part of repeated social infrastructure development. After 1970 the stock value of educational facilities rapidly increased, and in 1997 educational facilities accounted for 12.2% of all social infrastructures (with a gross stock value of 70 trillion yen) [5]. Furthermore, among educational facilities, the average serviceable life of schools and academic facilities is considered to be about 30 years, short in comparison to the average serviceable life of civil infrastructures. In fact, in 2003, the gross total area of elementary and junior high schools was 160.9 million square meters, of which 41.5% was 20- to 29-year-old facilities and 29.0% was 30 to 39 [5], the aging of which is beginning to be actualized. The results of repair and reconstruction cost estimations based on this data are shown in Fig.1. Simply because their average serviceable life is short, the repair and reconstruction costs of educational facilities are becoming greater than those of civil infrastructures. In conclusion, the asset management of educational facilities is a problem that has tremendous social urgency.

As points to consider in the asset management of educational facilities, when carrying out evaluations of the condition of the facility or components, one must consider not only structural safety, but also usability and convenience for users, and furthermore aesthetic aspect, all of which may be listed as important evaluation factors. In other words, they characteristically have a large number of components that users will come into direct contact with and components that users will view directly. For example, in the case of doors and window frames, even though they do not in any way influence safety, they will be targeted for repair or reconstruction if they fail to open and close and seem to damage usability for users. Furthermore, if there is partial damage to the tiles of exterior walls or partial deterioration of paint, repairs may be carried out for aesthetic reasons. Therefore, in the asset management of educational facilities, even when the authors speak of deterioration predictions a simple physical deterioration prediction targeting structural safety is not sufficient, and rather a general performance prediction that includes structural safety considerations is necessary. At the present time, other than visual inspection, this kind of general performance evaluation does not exist, and a statistical deterioration prediction method (statistical performance prediction method) based upon visual inspection data would be effective.

2.2 Hazard model and the heterogeneity of the deterioration process

In traditional hazards analysis, it is assumed that the target facility is entirely built of the same material, with the aim of modeling deterioration phenomena that arrive randomly in accordance with certain hazard functions. In hazard analysis,

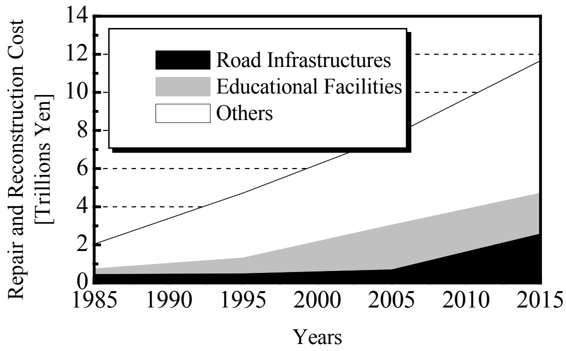

Fig.1 Transitions in the repair and reconstruction costs of infrastructures [5]

the occurrence process of random deterioration phenomena is modeled, and the hazard function, a deterministic probability model, is used. However, in large-scale facilities, such as the educational facilities targeted in the empirical study of this paper, it is not necessarily possible to express the hazard rate of each individual facility component with the same hazard rate. Rather, it is more natural to consider that the hazard rate for each type of component will have a different respective hazard rate. For the management and operation of large-scale facilities, the consideration of repair and reconstruction plans for these many components is a critical issue. In this way, as a method that expresses the heterogeneity of a hazard rate that considers the differences in the component types, we can consider 1) a method in which differences in component properties are expressed as dummy variables and deterioration estimation is carried out, and 2) a method in which it is assumed that the hazard rate will be subject to a particular probability distribution for each component group after which a deterioration estimation is carried out. The fist method has the advantage of being simple and easy to understand. On the other hand, it is problematic because as the number of components increases the number of dummy variables (which express component properties) increases, and the estimation accuracy of the model decreases remarkably. In addition, an increase in explanatory variables is directly connected to an increase in field observation and inspection items, increasing the burdens in practice. Furthermore, because the heterogeneity of the deterioration process may be controlled by factors that are not possible to observe, it is essential that a more efficient deterioration prediction method be developed. Educational facilities are constructed from an extremely large number of components, and estimating a hazard model that makes use of dummy variables is not practical. In order to express the heterogeneity of the deterioration process, there are limits to the refinement of a Weibull hazard model by increasing explanatory variables and so on. As long as innate facility information is expressed as explanatory variables, the estimate accuracy and efficiency will inevitable decrease. Therefore, in this study the authors have employed a mixed hazard model in which, depending on the type of component, the heterogeneity of the hazard rate is expressed as a probability distribution to model the deterioration process of a facility.

Research into hazard analyses that consider the heterogeneity of hazard rates is accumulating. In particular, there is a large accumulation of studies regarding mixed hazard models in which there exists a heterogeneous hazard rate for each individual sample. In mixed hazard models, it is considered that heterogeneity parameters controlling the hazard function are distributed with being subject to a probability density function. In addition, a hazard function is defined by probabilistic convolutions of the probability distribution of the hazard function and heterogeneous parameters. In regards to a mixed hazard model, Kaito et al. [6] have modeled the arrival process of road obstacles and made a case study of the application to asset management. On the other hand, educational facilities are composed of a large number of component types, such as roofs, exterior walls, doors and eaves. To put it another way, it can be anticipated that there exist component groups that require homogenous hazard rates, and that each group's hazard rate has an inherent probability function. In this study, it is considered that a mixed hazard model in which these probability error items are assumed to have a gamma distribution.

2.3. The basic/fundamental frame of the model

Suppose the modeling of the facility deterioration process. The case in which it is composed of N types of components is considered. Component type i ($i = 1,\ldots, N$) is composed of N_i samples. At time $t = 0$, a continuous-time axis that continues infinitely is introduced. The facility is considered to be completely repaired at time $t = 0$. From time $t = 0$, the deterioration of each component begins to progress. Should the deterioration of a component progresses until an unallowable level, that component will be repaired immediately. It is assumed that a component that has been repaired possesses the same deterioration performance of the new constructed ones. Now, from among the components, take a look at component type 1, for which the authors have information at time $t = T$, when T amount of time has passed from time $t = 0$. Component type 1 is composed of N_1 components. Among these, component 1_A is considered to have not been determined to be deteriorated even once from time $t = 0$ yet. The observed time of use of component 1_A is T, and the lifetime of the concerned component can be understood to be at least longer than time of use T. On the other hand, component 1_B is considered to have already deteriorated at times T_1 and T_2. Until the first deterioration time point, the facility lifespan is $\zeta = T_1$, and until the second deterioration time point the lifespan is $\zeta = T_2 - T_1$.

Here, suppose that each component has time-dependent deterioration properties. For the time-dependent type components, as indicated in Fig.2, the more time that has elapsed since the most recent repair time, the higher the occurrence probability of deterioration. Assume that the lifespan distribution of components characterized by this kind of the time-dependent deterioration types follows a Weibull distribution. Furthermore, suppose that the hazard function of different components can be expressed as indicated in Fig.2. With respect to the benchmark hazard

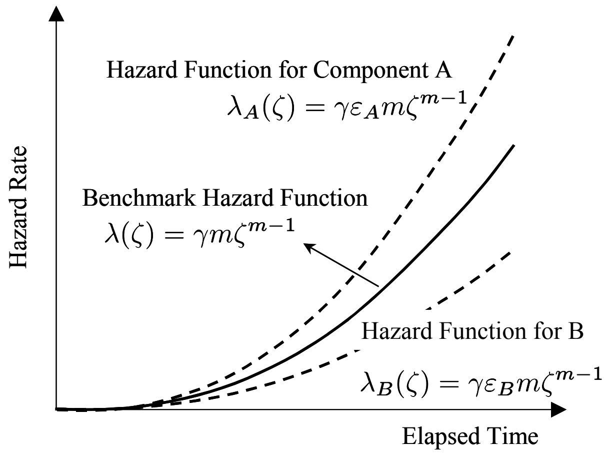

Fig.2 The heterogeneity of the hazard function

function, these hazard functions have a relationship of being mutually amplified a constant number of times, or being decreased. In this way, a model that expresses hazard function groups that have a relationship of mutual analogous amplification or decrease can be called a proportional hazards model [2]. When the deterioration process of components that compose each type can be expressed by a random hazard model, the heterogeneity of the hazard model can be expressed by the probability distribution of the proportional constant of the hazard rate. Educational facilities are composed of a large number of types, but it is not uncommon for the number of individual components found in each type to not be so large. By estimating the benchmark Weibull hazard function parameters and the probability function parameters, which express the heterogeneous distribution of the proportionality constant between types, the random proportional Weibull deterioration hazards model can easily express heterogeneity between types. For this reason, the random proportional Weibull hazard model proposed in this study is ideal to express the deterioration process of educational facilities, which are made of a small number of a wide variety of components. However, in order to employ the random proportional Weibull hazard model, it is presupposed that the heterogeneity of the hazard rate between types has a mutual proportional relationship. Hypothesis testing related to the existence of a proportional relationship is mentioned again in section **4.2**.

3. RANDOM PROPORTIONAL HAZARD MODEL

3.1. Formulation of Weibull hazard model

The random proportional Weibull hazard model is a Weibull hazard model [3] that considers the heterogeneity of the hazard rate between components. The details of hazard models in general may be found in the references [7-8]. Here, before formulating the random proportional Weibull hazards model, a basic concept of the Weibull hazard model is given for the convenience of readers.

Now, certain educational facilities have begun service, and the authors focus on the period up to when the condition of the components reaches the limit of service (hereafter called the lifespan). During this time, these components are in continuous service, and whether or not they have reached the limit of service is evaluated in periodical inspections. It is assumed that the lifespan of components is expressed by the random variable ζ, and distributed according to the probability density function $f(\zeta)$ and cumulative distribution function $F(\zeta)$. However, the domain of the lifespan ζ is defined as [0, ∞]. Now, the survival probability $\tilde{F}(\zeta)$ of a component from the initial time point to the arbitrary time point $\zeta \in [0,\infty]$ is the value subtracting the component failure probability $F(\zeta)$ from the whole occurrence probability 1, and defined

$$\tilde{F}(\zeta) = 1 - F(\zeta) \tag{1}$$

Here, the conditional probability that the component survives until time ζ, and moreover fails (progresses its deterioration until an unacceptable level) during the time interval period [ζ, $\zeta+\Delta\zeta$] for the first time can be expressed as

$$\lambda(\zeta)\Delta\zeta = \frac{f(\zeta)\Delta\zeta}{\tilde{F}(\zeta)} \tag{2}$$

At this time, the probability density $\lambda(\zeta)$ is called the "hazard function." By differentiating both sides of equation (1) with respect to ζ, the following equation is obtained.

$$\frac{d\tilde{F}(\zeta)}{d\zeta} = -f(\zeta) \tag{3}$$

Using equation (3), equation (2) can be expressed as

$$\lambda(\zeta) = \frac{d}{d\zeta}\left(-\log\tilde{F}(\zeta)\right) \tag{4}$$

Here, considering that $\tilde{F}(0) = 1 - F(0) = 1$, if equation (4) is integrated,

$$\int_0^{\zeta} \lambda(u)du = \left[-\log\tilde{F}(u)\right]_0^{\zeta} \tag{5}$$

is obtained. Consequently, if the hazard function $\lambda(\zeta)$ is used, the cumulative survival probability $\tilde{F}(\zeta)$ of a component until time point ζ is represented as

$$\tilde{F}(\zeta) = \exp\left[-\int_0^{\zeta}\lambda(u)du\right] \tag{6}$$

In above equation, if the function form of the hazard function $\lambda(\zeta)$ is determined, the component survival probability $\tilde{F}(\zeta)$ can be derived. Furthermore, from the relationship of equation (1), the cumulative deterioration probability $F(\zeta)$ of components is obtained. Here, as the hazard function, consider the Weibull hazard function

$$\lambda(\zeta) = \gamma m \zeta^{m-1} \tag{7}$$

where γ is a parameter that expresses the arrival density and m is an acceleration parameter that represents the tendency of the hazard rate to increase over time. When the Weibull hazard function is used, the component lifespan probability density function $f(\zeta)$, and the component survival probability $\tilde{F}(\zeta)$ become as follows respectively.

$$f(\zeta) = \gamma m \zeta^{m-1}\exp(-\gamma\zeta^m) \tag{8a}$$

$$\tilde{F}(\zeta) = \exp(-\gamma\zeta^m) \tag{8b}$$

3.2. Random proportional Weibull hazard model

A certain component of a facility is considered to be classified into N kinds of component types. A total of N_i of the ith (i = 1,…,N) component type exist. Furthermore, focus on the jth (j = 1,…,N_i) component of type i. The time that has elapsed since the aforementioned component was reconstructed is represented by ζ_i^j. Suppose that the arrival rate of deterioration events for each component conforms to a Weibull deterioration hazard function

$$\lambda_i(\zeta_i^j) = \gamma m \varepsilon_i (\zeta_i^j)^{m-1} \tag{9}$$

In equation (9), the parameter ε_i (called the heterogeneity parameter below), which represents the heterogeneity of the hazard rate of type i, has been added to the Weibull hazard function from equation (7). The heterogeneity parameter takes a common value for components of the same type. However, when component type differs, it takes different values. In actuality, the heterogeneity parameter takes on a deterministic value, but is an impossible parameter for an observer to observe. In addition, the lifespan probability density function $f_i(\zeta_i^j)$ of type i component j, and the survival probability $\tilde{F}(\zeta_i^j)$ are respectively expressed as

$$f_i(\zeta_i^j) = \gamma m \varepsilon_i (\zeta_i^j)^{m-1}\exp\left\{-\gamma\varepsilon_i(\zeta_i^j)^m\right\} \tag{10a}$$

$$\tilde{F}(\zeta_i^j) = \exp\left\{-\gamma\varepsilon_i(\zeta_i^j)^m\right\} \tag{10b}$$

Now, the value of the heterogeneity parameter is one of the observations from random variables that cannot be directly measured by the observer, but it is known to be distributed in accordance with the probability density function $g(\varepsilon)$. That is the Weibull hazard model (9) has an identical deterioration acceleration parameter m for all types of components, but for each component the arrival ratio $\gamma m \varepsilon_i$ differs proportionally, and the individuality of deterioration is expressed. Regarding hypothesis testing of the homogeneity (below, proportionality) of the acceleration parameter, we make another investigation. In this study, for each targeted component, a Weibull hazard model in which the hazard arrival ratio is a observation from a probability distribution is called a random proportional Weibull hazard [4].

Here, suppose that the probability distribution of the heterogeneity parameter conforms to a Gamma distribution. The Gamma distribution, as a special form, includes the exponential distribution, and has the advantage that it can express the exponential family probability distribution function that is defined on the interval $[0,\infty]$. Here, suppose that the parameter γ represents the average hazard arrival ratio between types, and the heterogeneity parameter ε_i is a observation from a Gamma distribution with average 1 and variance ϕ^{-1} and is a probabilistic error term. The Gamma function is defined on the interval $[0,\infty]$, and with respect to an arbitrary explanatory variable and probabilistic error term, the right side of equation (9) is assured to take the positive value. In general, the probability density function $g(\varepsilon:\alpha,\beta)$ of the Gamma distribution $G(\alpha,\beta)$ can be defined as

$$g(\varepsilon:\alpha,\beta)=\frac{1}{\beta^{\alpha}\mathbf{\Gamma}(\alpha)}\varepsilon^{\alpha-1}\exp\left(-\frac{\varepsilon}{\beta}\right) \tag{11}$$

The average of the Gamma distribution $G(\alpha,\beta)$ is $\mu=\alpha\beta$, and the variance $\sigma^2=\alpha\beta^2$. In addition, $\mathbf{\Gamma}(\,)$ is a Gamma function. Furthermore, the probability density function $\bar{g}(\varepsilon:\phi)$ of the standard Gamma distribution that has an average 1 and variance ϕ^{-1} is expressed as

$$\bar{g}(\varepsilon:\phi)=\frac{\phi^{\phi}}{\mathbf{\Gamma}(\phi)}\varepsilon^{\phi-1}\exp(-\phi\varepsilon) \tag{12}$$

3.3. Two Steps estimation method for the model

In a random proportional Weibull hazard model, a total of 3+N unknown parameters exist, the arrival density parameter γ, acceleration parameter m, heterogeneity parameter ε_i (i = 1,…, N), which differs for each component, and the distribution parameter ϕ of the heterogeneity parameter. In the case of an ordinary Weibull hazard model, it is enough to estimate the parameters γ and m from deterioration data record. However, in the random proportional Weibull hazard model, besides these two parameters, it is necessary to pursue the probability distribution parameter ϕ of the heterogeneity parameter and the heterogeneity parameter ε_i (i = 1,…, N) for each component type.

Now, let us suppose that the deterioration history database of the facility is available. The database contains information relating to the time of deterioration (repair) of all components from the time the targeted components began service. Express the deterioration record of components as $\mathbf{\Xi}=(\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_N)$, where $\xi_i=\{(\delta_i^1,\zeta_i^1),\cdots,(\delta_i^{N_i},\zeta_i^{N_i})\}$ (i = 1,…, N). In addition, δ_i^j is a dummy variable that takes the value 1 if type i component j (j = 1,…, N) has deteriorated, and takes the value 0 if it has not deteriorated, and ζ_i^j is the period of service of type i component j, that is, when $\delta_i^j=0$, ζ_i^j means the length of time from the previous repair or reconstruction to the present. On the other hand, when $\delta_i^j=1$, ζ_i^j indicates lifespan. Here, suppose that the heterogeneity parameter ε_i is given. At this time, the conditional likelihood $\ell_i(\boldsymbol{\xi}_i:\gamma,m,\varepsilon_i)$ for the observed data ξ_i for type i is expressed as

$$\ell_i(\boldsymbol{\xi}_i:\gamma,m,\varepsilon_i)=\prod_{j=1}^{N_i}\left\{\widetilde{F}(\zeta_i^j:\gamma,m,\varepsilon_i)\right\}^{(1-\delta_i^j)} \cdot\left\{f_i(\zeta_i^j:\gamma,m,\varepsilon_i)\right\}^{\delta_i^j} \tag{13}$$

However, in the above equation, it is explicitly indicated that the lifespan probability density function $f(\zeta_i^j:\gamma,m,\varepsilon_i)$ and the survival function $\widetilde{F}(\zeta_i^j:\gamma,m,\varepsilon_i)$ are described as functions of parameters γ, m and ε_i. Here, if the heterogeneity ε_i is distributed according to the standard Gamma distribution $\bar{g}(\varepsilon_i:\phi)$, the likelihood function for the observation data ξ_i is

$$\begin{aligned}L_i(\boldsymbol{\xi}_i:\boldsymbol{\theta})&=\int_0^{\infty}\prod_{j=1}^{N_i}\left\{\widetilde{F}(\zeta_i^j:\gamma,m,\varepsilon_i)\right\}^{(1-\delta_i^j)}\\&\cdot\left\{f_i(\zeta_i^j:\gamma,m,\varepsilon_i)\right\}^{\delta_i^j}\bar{g}(\varepsilon_i:\phi)d\varepsilon_i\\&=\frac{\phi^{\phi}}{\mathbf{\Gamma}(\phi)}\prod_{j=1}^{N_i}\left\{\gamma m(\zeta_i^j)^{m-1}\right\}^{\delta_i^j}\\&\cdot\int_0^{\infty}\varepsilon_i^{s_i+\phi-1}\exp\{-(\phi+\gamma\tau_i)\varepsilon_i\}d\varepsilon_i\end{aligned} \tag{14}$$

However, $s_i=\sum_{j=1}^{N_i}\delta_i^j$ and $\tau_i=\sum_{j=1}^{N_i}(\zeta_i^j)^m$. In the above equation, with respect to all type i components, the heterogeneity parameter ε_i takes a common value. To express this, it should be noted that the likelihood function $L_i(\boldsymbol{\xi}_i:\boldsymbol{\theta})$ is defined as an expected value related to the probability function ε_i of the conditional likelihood $\ell_i(\boldsymbol{\xi}_i:\gamma,m,\varepsilon_i)$. Here, if a variable transformation $x_i=\varepsilon_i(\phi+\gamma\tau_i)$ is carried out

$$\begin{aligned}L_i(\boldsymbol{\xi}_i:\boldsymbol{\theta})&=\frac{\phi^{\phi}}{\mathbf{\Gamma}(\phi)}\prod_{j=1}^{N_i}\left\{\gamma m(\zeta_i^j)^{m-1}\right\}^{\delta_i^j}\\&\cdot\int_0^{\infty}\left(\frac{x_i}{\phi+\gamma\tau_i}\right)^{s_i+\phi-1}\exp\{-x_i\}\frac{dx_i}{\phi+\gamma\tau_i}\\&=\frac{\phi^{\phi}}{(\phi+\gamma\tau_i)^{s_i+\phi}}\frac{\mathbf{\Gamma}(s_i+\phi)}{\mathbf{\Gamma}(\phi)}\cdot\prod_{j=1}^{N_i}\left\{\gamma m(\zeta_i^j)^{m-1}\right\}^{\delta_i^j}\end{aligned} \tag{15}$$

is obtained. Therefore, the logarithmic likelihood function for the observed data $\mathbf{\Xi}=(\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_N)$ can be expressed as

$$\begin{aligned}\ln L(\mathbf{\Xi},\boldsymbol{\theta})&=\sum_{i=1}^{N}\ln L_i(\boldsymbol{\xi}_i:\boldsymbol{\theta})\\&=\sum_{i=1}^{N}\Big[\phi\ln\phi-(s_i+\phi)\ln(\phi+\gamma\tau_i)+\ln\mathbf{\Gamma}(s_i+\phi)\\&-\ln\mathbf{\Gamma}(\phi)+\sum_{j=1}^{N_a}\delta_i^j\left\{\ln\gamma+\ln m+(m-1)\ln\zeta_i^j\right\}\Big]\end{aligned} \tag{16}$$

However, each element of $\boldsymbol{\theta}=(\theta_1,\theta_2,\theta_3)$ is expressed as $\theta_1=\gamma$, $\theta_2=m$ and $\theta_3=\phi$. The maximum likelihood estimator of the parameter $\boldsymbol{\theta}$ that maximizes logarithmic likelihood function (16) can be given as $\hat{\boldsymbol{\theta}}=(\hat{\theta}_1,\hat{\theta}_2,\hat{\theta}_3)$, which simultaneously satisfies

$$\frac{\partial\ln L(\hat{\boldsymbol{\theta}},\mathbf{\Xi})}{\partial\theta_i}=0 \tag{17}$$

Furthermore, the estimator $\hat{\mathbf{\Sigma}}(\hat{\boldsymbol{\theta}})$ of the asymptotic covariance matrix can be expressed as

$$\hat{\mathbf{\Sigma}}(\hat{\boldsymbol{\theta}})=\left[\frac{\partial^2\ln L(\hat{\boldsymbol{\theta}},\mathbf{\Xi})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right]^{-1} \tag{18}$$

However, the inverse matrix of the right side of the above formula is the inverse matrix of a 3 x 3 Fisher information matrix that consists of elements $\partial^2\ln L(\hat{\boldsymbol{\theta}},\mathbf{\Xi})/\partial\theta_i\partial\theta_j$. The maximum likelihood estimator of the parameter is obtained by solving the three dimensional nonlinear simultaneous equation (17). In this study, the maximum likelihood

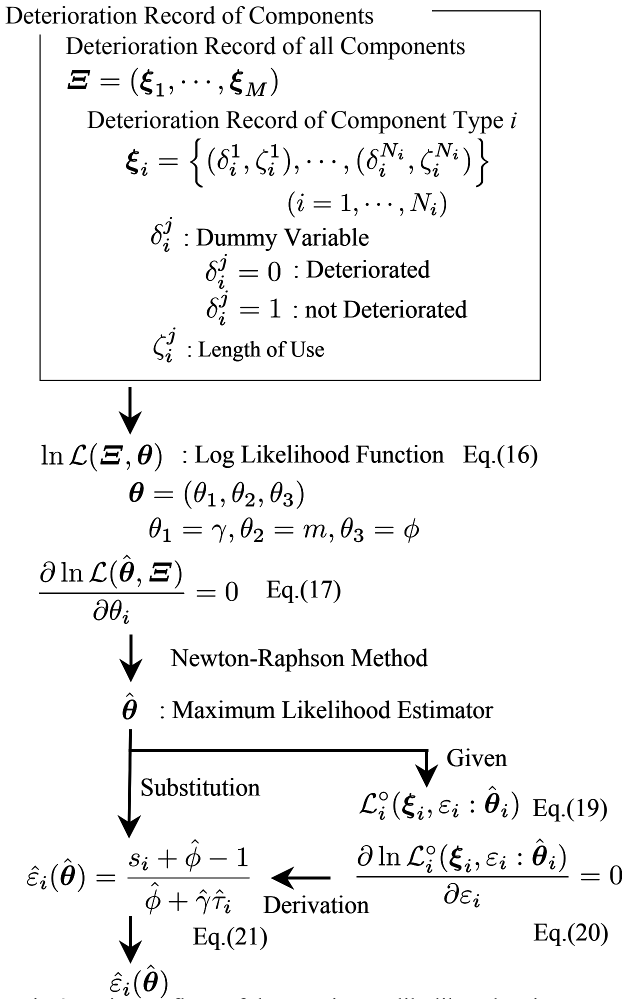

Fig.3 Estimate flow of the maximum likelihood estimator

estimator was obtained using the Newton-Raphson Method. If maximum likelihood estimator $\hat{\boldsymbol{\theta}}$ is obtained, using a covariance matrix estimator $\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})$, *t*-test statistic can be also estimated.

Next, with the parameter vector's maximum likelihood estimator $\hat{\boldsymbol{\theta}}$ as a given, the maximum likelihood estimator of the heterogeneity parameter ε_i ($i = 1,\ldots, N$) is obtained. Here, the partial likelihood function is defined as

$$L_i^0(\boldsymbol{\xi}_i, \varepsilon_i : \hat{\boldsymbol{\theta}}) = \frac{\hat{\phi}^{\hat{\phi}}}{\boldsymbol{\Gamma}(\hat{\phi})} \prod_{j=1}^{N_i} \left\{\hat{\gamma}\hat{m}(\zeta_i^j)^{\hat{m}-1}\right\}^{\delta_i^j} \varepsilon_i^{s_i+\hat{\phi}-1} \cdot \exp\left\{-(\hat{\phi} + \hat{\gamma}\hat{\tau}_i)\varepsilon_i\right\} \tag{19}$$

Here, $\hat{\tau}_i = \sum_{j=1}^{N_i} (\zeta_i^j)^{\hat{m}}$. At this time, the maximum likelihood estimator of the heterogeneity parameter ε_i ($i = 1,\ldots, N$) can be obtained as $\hat{\varepsilon}_i^o$ that satisfies

$$\frac{\partial \ln L_i^0(\boldsymbol{\xi}_i, \varepsilon_i : \hat{\boldsymbol{\theta}})}{\partial \varepsilon_i} = 0 \tag{20}$$

The maximum likelihood estimator of the heterogeneity parameter obtained in this way is an estimator that was obtained with the given parameter $\hat{\boldsymbol{\theta}} = (\hat{\gamma}, \hat{m}, \hat{\phi})$. In order to clearly describe this, the solution of equation (20) is expressed as $\hat{\varepsilon}_i(\hat{\boldsymbol{\theta}})$. From equations (19) and (20), if $\hat{\varepsilon}_i(\hat{\boldsymbol{\theta}})$ is specifically estimated, the following equation is obtained:

$$\hat{\varepsilon}_i(\hat{\boldsymbol{\theta}}) = \frac{s_i + \hat{\phi} - 1}{\hat{\phi} + \hat{\gamma}\hat{\tau}_i} \tag{21}$$

The above two-step maximum likelihood estimator estimation flow is shown in Fig.3.

4. AN EMPIRICAL STUDY

4.1. Overview of the case of application

A random proportional Weibull hazard model estimation is attempted for a certain university facility. The condition state of this facility is accumulated through visual inspections. The inspection period is three years. The condition state of the facility is evaluated as either possible to use (○) or not possible to use (×). This facility group has all been classified into 33 regions and is located in each one. The oldest facility was built 73 years ago. This time, the data used in estimations was the most recent visual inspection data, which was collected in 2006. Below, the authors will use the technical term deterioration, but as mentioned before, the visual inspection data used in this study as deterioration is defined not only as physical damage to components, but is also as loss of pleasantness and convenience of the facility that is judged to need repair.

For the specific estimate target of the random proportional Weibull hazard model, exterior wall components, for which the greatest abundance of data was obtained, were focused on. Exterior walls can be classified into five types: tile, multi-layer finish painted, thin finish painted, metal, and concrete blocks. In addition, because there are a large number of exterior walls, an extremely small number of components exist for which, due to an initial failure, the time period from the start of service until the deterioration time point was remarkable short. For this reason, in this estimate, exterior walls for which repair was carried out within one year from start of service are deemed to be initial failure samples, and such samples were excluded in advance. After the above preliminary preparation, the sum total by type of exterior walls samples that could be used in the estimate were: 77 tile samples, 35 multi-layered finish painted samples, 52 thin finish painted samples, 20 metal samples, and 26 concrete blocks samples. Therefore, the total number of exterior wall samples was 210.

4.2. Proportionality assumption testing

In random proportional Weibull hazard models, a proportionality assumption is made in which all types of components have an identical acceleration parameter $\hat{m}$. Therefore, the differences in the deterioration process between exterior wall types (tile, multi-layered finish painted, thin finish painted, metal and concrete block) can be considered to be aggregated in the heterogeneity parameter. Based on actual data, the authors propose an assumption testing method to determine whether or not the to an arbitrary explanatory variable and probabilistic error term, the right side of equation (9) is assured to take the positive value. In general, the probability density function $g(\varepsilon:\alpha,\beta)$ of the Gamma distribution $G(\alpha,\beta)$ can be defined as

$$g(\varepsilon:\alpha,\beta)=\frac{1}{\beta^{\alpha}\mathbf{\Gamma}(\alpha)}\varepsilon^{\alpha-1}\exp\left(-\frac{\varepsilon}{\beta}\right) \tag{11}$$

The average of the Gamma distribution $G(\alpha,\beta)$ is $\mu=\alpha\beta$, and the variance $\sigma^2=\alpha\beta^2$. In addition, $\mathbf{\Gamma}(\,)$ is a Gamma function. Furthermore, the probability density function $\overline{g}(\varepsilon:\phi)$ of the standard Gamma distribution that has an average 1 and variance ϕ^{-1} is expressed as

$$\overline{g}(\varepsilon:\phi)=\frac{\phi^{\phi}}{\mathbf{\Gamma}(\phi)}\varepsilon^{\phi-1}\exp(-\phi\varepsilon) \tag{12}$$

3.3. Two Steps estimation method for the model

In a random proportional Weibull hazard model, a total of 3+*N* unknown parameters exist, the arrival density parameter γ, acceleration parameter *m*, heterogeneity parameter ε_i (i = 1,…, *N*), which differs for each component, and the distribution parameter ϕ of the heterogeneity parameter. In the case of an ordinary Weibull hazard model, it is enough to estimate the parameters γ and m from deterioration data record. However, in the random proportional Weibull hazard model, besides these two parameters, it is necessary to pursue the probability distribution parameter ϕ of the heterogeneity parameter and the heterogeneity parameter ε_i (i = 1,…, *N*) for each component type.

Now, let us suppose that the deterioration history database of the facility is available. The database contains information relating to the time of deterioration (repair) of all components from the time the targeted components began service. Express the deterioration record of components as $\mathbf{\Xi}=(\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_N)$, where $\xi_i=\left\{(\delta_i^1,\zeta_i^1),\cdots,(\delta_i^{N_i},\zeta_i^{N_i})\right\}$ (i = 1,…, *N*). In addition, δ_i^j is a dummy variable that takes the value 1 if type *i* component *j* (*j* = 1,…, *N*) has deteriorated, and takes the value 0 if it has not deteriorated, and ζ_i^j is the period of service of type *i* component *j*, that is, when $\delta_i^j=0$, ζ_i^j means the length of time from the previous repair or reconstruction to the present. On the other hand, when $\delta_i^j=1$, ζ_i^j indicates lifespan. Here, suppose that the heterogeneity parameter ε_i is given. At this time, the conditional likelihood $\ell_i(\boldsymbol{\xi}_i:\gamma,m,\varepsilon_i)$ for the observed data ξ_i for type *i* is expressed as

$$\begin{aligned}\ell_i(\boldsymbol{\xi}_i:\gamma,m,\varepsilon_i)=&\prod_{j=1}^{N_i}\left\{\widetilde{F}(\zeta_i^j:\gamma,m,\varepsilon_i)\right\}^{(1-\delta_i^j)}\\&\cdot\left\{f_i(\zeta_i^j:\gamma,m,\varepsilon_i)\right\}^{\delta_i^j}\end{aligned} \tag{13}$$

However, in the above equation, it is explicitly indicated that the lifespan probability density function $f(\zeta_i^j:\gamma,m,\varepsilon_i)$ and the survival function $\widetilde{F}(\zeta_i^j:\gamma,m,\varepsilon_i)$ are described as functions of parameters γ, m and ε_i. Here, if the heterogeneity ε_i is distributed according to the standard Gamma distribution $\overline{g}(\varepsilon_i:\phi)$, the likelihood function for the observation data ξ_i is

$$\begin{aligned}L_i(\boldsymbol{\xi}_i:\boldsymbol{\theta})=&\int_0^{\infty}\prod_{j=1}^{N_i}\left\{\widetilde{F}(\zeta_i^j:\gamma,m,\varepsilon_i)\right\}^{(1-\delta_i^j)}\\&\cdot\left\{f_i(\zeta_i^j:\gamma,m,\varepsilon_i)\right\}^{\delta_i^j}\overline{g}(\varepsilon_i:\phi)d\varepsilon_i\\=&\frac{\phi^{\phi}}{\mathbf{\Gamma}(\phi)}\prod_{j=1}^{N_i}\left\{\gamma m(\zeta_i^j)^{m-1}\right\}^{\delta_i^j}\\&\cdot\int_0^{\infty}\varepsilon_i^{s_i+\phi-1}\exp\{-(\phi+\gamma\tau_i)\varepsilon_i\}d\varepsilon_i\end{aligned} \tag{14}$$

However, $s_i=\sum_{j=1}^{N_i}\delta_i^j$ and $\tau_i=\sum_{j=1}^{N_i}(\zeta_i^j)^m$. In the above equation, with respect to all type i components, the heterogeneity parameter ε_i takes a common value. To express this, it should be noted that the likelihood function $L_i(\boldsymbol{\xi}_i:\boldsymbol{\theta})$ is defined as an expected value related to the probability function ε_i of the conditional likelihood $\ell_i(\boldsymbol{\xi}_i:\gamma,m,\varepsilon_i)$. Here, if a variable transformation $x_i=\varepsilon_i(\phi+\gamma\tau_i)$ is carried out

$$\begin{aligned}L_i(\boldsymbol{\xi}_i:\boldsymbol{\theta})=&\frac{\phi^{\phi}}{\mathbf{\Gamma}(\phi)}\prod_{j=1}^{N_i}\left\{\gamma m(\zeta_i^j)^{m-1}\right\}^{\delta_i^j}\\&\cdot\int_0^{\infty}\left(\frac{x_i}{\phi+\gamma\tau_i}\right)^{s_i+\phi-1}\exp\{-x_i\}\frac{dx_i}{\phi+\gamma\tau_i}\\=&\frac{\phi^{\phi}}{(\phi+\gamma\tau_i)^{s_i+\phi}}\frac{\mathbf{\Gamma}(s_i+\phi)}{\mathbf{\Gamma}(\phi)}\cdot\prod_{j=1}^{N_i}\left\{\gamma m(\zeta_i^j)^{m-1}\right\}^{\delta_i^j}\end{aligned} \tag{15}$$

is obtained. Therefore, the logarithmic likelihood function for the observed data $\mathbf{\Xi}=(\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_N)$ can be expressed as

$$\begin{aligned}\ln L(\mathbf{\Xi},\boldsymbol{\theta})=&\sum_{i=1}^{N}\ln L_i(\boldsymbol{\xi}_i:\boldsymbol{\theta})\\=&\sum_{i=1}^{N}\Big[\phi\ln\phi-(s_i+\phi)\ln(\phi+\gamma\tau_i)+\ln\mathbf{\Gamma}(s_i+\phi)\\&-\ln\mathbf{\Gamma}(\phi)+\sum_{j=1}^{N_a}\delta_i^j\left\{\ln\gamma+\ln m+(m-1)\ln\zeta_i^j\right\}\Big]\end{aligned} \tag{16}$$

However, each element of $\boldsymbol{\theta}=(\theta_1,\theta_2,\theta_3)$ is expressed as $\theta_1=\gamma$, $\theta_2=m$ and $\theta_3=\phi$. The maximum likelihood estimator of the parameter $\boldsymbol{\theta}$ that maximizes logarithmic likelihood function (16) can be given as $\hat{\boldsymbol{\theta}}=(\hat{\theta}_1,\hat{\theta}_2,\hat{\theta}_3)$, which simultaneously satisfies

$$\frac{\partial\ln L(\hat{\boldsymbol{\theta}},\mathbf{\Xi})}{\partial\theta_i}=0 \tag{17}$$

Furthermore, the estimator $\hat{\mathbf{\Sigma}}(\hat{\boldsymbol{\theta}})$ of the asymptotic covariance matrix can be expressed as

$$\hat{\mathbf{\Sigma}}(\hat{\boldsymbol{\theta}})=\left[\frac{\partial^2\ln L(\hat{\boldsymbol{\theta}},\mathbf{\Xi})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right]^{-1} \tag{18}$$

However, the inverse matrix of the right side of the above formula is the inverse matrix of a 3 x 3 Fisher information matrix that consists of elements $\partial^2\ln L(\hat{\boldsymbol{\theta}},\mathbf{\Xi})/\partial\theta_i\partial\theta_j$. The maximum likelihood estimator of the parameter is obtained by solving the three dimensional nonlinear simultaneous equation (17). In this study, the maximum likelihood

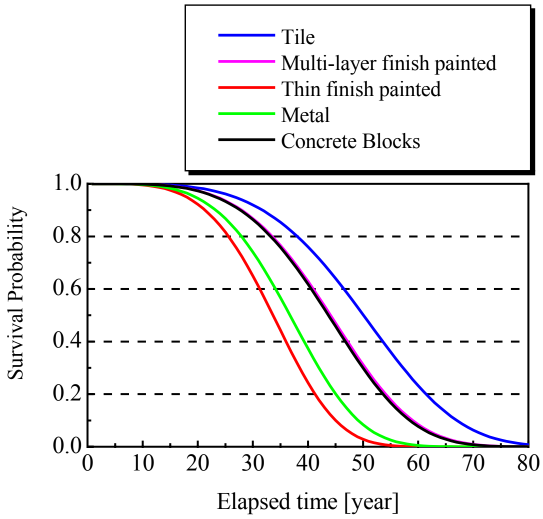

Fig.4 Survival Probability of Exterior Walls

multi-layer finish painted, 33.7 years for thin finish painted, 37.0 years for metal, and 43.7 years for concrete blocks, and as mentioned before, the deterioration progress of thin finish painted is the fastest and that of tiles is the slowest. When thin finish painted has been in use for 25 years, its survival probability is about 81.3%, when in use for 50 years its survival probability is about 0.03%. From this, we can see that for thin finish painted, as the period of service becomes longer, the deterioration probability becomes larger at an accelerating pace. In addition, when tile has been in use for 25 years, its survival probability is about 96.5%, and when in use for 50 years its survival probability is about 50.3%, and in the same way as multi-layer finish painted, metal, and concrete blocks, as the period of service becomes longer, a tendency for the deterioration probability to become larger at an accelerating pace can be confirmed.

From the above, it can be understood that even for the same exterior wall component, for each type there is wide variation in the heterogeneity parameter. More specifically, to estimate a hazard model for a large-scale facility such as an educational facility, which is made up of various kinds of components, it can be said that a random proportional Weibull hazard function that uses a mixed distribution is effective. In addition, the deterioration survival probability can be estimated for individual components, which could not be considered if the hazard rate were simply treated deterministically. Therefore, it is possible to expect this to contribute to the refinement of asset management.

5. CONCLUSIONS

In this study, a deterioration prediction for component groups that make up educational facilities was implemented. In so doing, the authors focused on each component being made up of a small number of a variety of types, and pointed out that a hazard model that could express the heterogeneity of the hazard rate between types would be needed. In this way, to operationally express the heterogeneity of the hazard rate, a Weibull hazard model was used as a base model, and a random proportional Weibull hazard model was formulated in which the proportional heterogeneity of the hazard rate was expressed as a gamma function. Furthermore, using an application case that targeted an actual university facility, the effectiveness of the proposed hazard model was positively verified. In addition, in the application of the random proportional Weibull hazard model proposed in this study to asset management, there remain a number of issues. First, a model that uses the hazard model proposed in this study to estimate a multi-step deterioration process needs to be constructed. In general, the cost of repairs to infrastructures depends on the condition state of deterioration progress. For that reason, a variety of alternative repair strategies exist. In view of this, when developing repair strategies to minimize repair costs, mixed hazard models that describe multi-step deterioration processes need to be extended. Second, an application that can be used for asset management needs to be developed.

A part of this research has been carried out on the Program of Promotion of Environmental Improvement to Enhance Young Researcher’s Independence, the Special Coordination Funds for Promoting Science and Technology, Japan Ministry of Education, Culture, Sports, Science and Technology.

REFERENCES

1 Tsuda, Y., Kaito, K., Aoki, K. and Kobayashi, K., Estimating Markovian Transition Probabilities for Bridge Deterioration Forecasting, *Journal of Structural Eng. /Earthquake Eng.*, *JSCE*, Vol.23, No.2, pp.241s-256s, 2006.
2 Mishalani, R. and Madanat, S., Computation of Infrastructure Transition Probabilities using Stochastic duration Model, *ASCE Journal of Infrastructure Systems,* Vol.8, No.4, 2002.
3 Aoki, K., Yamamoto, K. and Kobayashi, K.: Estimating Hazard Models for Deterioration Forecasting, *Journal of Construction Management and Engineering*, *JSCE*, No.791/VI-67, pp.111-124, 2005 (in Japanese).
4 Kaito, K., Yamamoto, K., Obama, K., Okada, K. and Kobayashi, K., Random Proportional Weibull Hazard Model: An Application to Traffic Control Systems, *Journal of JSCE F* (submitted, in Japanese).
5 Cabinet Office, Director-General for Policy Planning, *Social Capital of Japan*, Printing Bureau, Ministry of Finance, 2002 (in Japanese).
6 Kaito, K., Kobayashi, K., Kato, T. and Ikuta, N., Road Patrol Frequency and Hazards Generations Risks, *Journal of JSCE F*, Vol.63, No.1, pp.16-34, 2007 (in Japanse).
7 Lancaster, T., *The Econometric Analysis of Transition Data*, Cambridge University Press, 1990.
8 Gourieroux, C., *Econometrics of Qualitative Dependent Variables*, Cambridge University Press, 2000.